LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-CONF-665482

Molecular Dynamics Simulations of the Antithrombin upon Ligand Recognition: A MM/GBSA Approach

X. Zhang, H. Perez-Sanchez, F. C. Lightstone

December 17, 2014

2015 International Work-Conference on Bioinformatics and Biomedical Engineering
Granada , Spain
April 15, 2015 through April 17, 2015

Disclaimer

Molecular Dynamics Simulations of the Antithrombin upon Ligand Recognition: A MM/GBSA Approach

Xiaohua Zhang[1], Horacio Perez-Sanchez[2]*, Felice C. Lightstone[1]*

1. Biosciences and Biotechnology Division, Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory (LLNL), Livermore, CA 94550
2. Bioinformatics and High Performance Computing Research Group, Department of Computer Science, Universidad Católica San Antonio de Murcia (UCAM), Spain

*To whom correspondence should be addressed: lightstone1@llnl.gov, hperez@ucam.edu

Abstract. The binding affinities of a series of antithrombin ligands have been calculated by using the MM/GBSA method. The calculated and experimental binding free energies have shown good correlation with a R^2 value of 0.69. Decomposition of the calculated binding free energy has revealed the electrostatic interactions in both solute and solvent play important role in determining the binding free energy. The increasing negative charge of compound provides more favorable electrostatic energy change but creates higher penalty for the solvation free energy. Such penalty is compensated by the electrostatic energy change, which results in better binding affinity. The best binder has the highest ligand efficiency.

Keywords: MM/GBSA, Molecular dynamics, Binding Affinity, Antithrombin, Heparin

1 Introduction

Antithrombin is a glycoprotein that plays crucial role in the regulation of blood coagulation by inactivating several enzymes of the coagulation system. It is an important drug target for the anticoagulant treatment. Antithrombin has two major isoforms, α and β, in the blood circulation[1]. α-antithrombin is the dominant form of antithrombin. It consists of 432 amino acids with 4 glycosylation sites, where an oligosaccharide occupying each glycosylation sites[2]. Heparin is the first compound that is identified and used as anticoagulant and antithrombotic agents. It is a sulfated polysaccharide containing a specific pentasaccharide fragment (Figure 1, NTP) that binds and activates the antithrombin[3]. This binding localized the function of antithrombin to inhibition of the serine proteases of the coagulation cascade in the bloodstream, which allows coagulant activity in damaged tissue outside the vascular system[2].

Due to increasing interests in clinical application, computational studies have been carried out to investigate the structure and behavior of antithrombin. Verli and co-workers performed molecular dynamics simulations to study the induce fit mechanism in antithrombin-heparin interaction and effects of glycosylation on heparin binding[4,5]. Several detailed conformational changes associated with heparin binding to antithrombin were revealed. They also confirmed an intermediate state between the native and activated forms of antithrombin. Because of the weak surface complementarity and the high charge density of the sulfated sugar chain, the docking of heparin to its protein partners presents a challenging task for computational docking. Wade and Bitomsky developed a protocol that can predict the heparin binding site correctly[6]. Navarro-Fernandez and colleagues screened a large database in silico and identified a new, non-polysaccharide scaffold able to interact with the heparin binding domain of antithrombin[7]. They predicted D-myo-inositol 3,4,5,6-tetrakisphosphate (Figure1, L1C4) to strongly interact with antithrombin, which was confirmed by experimental binding affinity study.

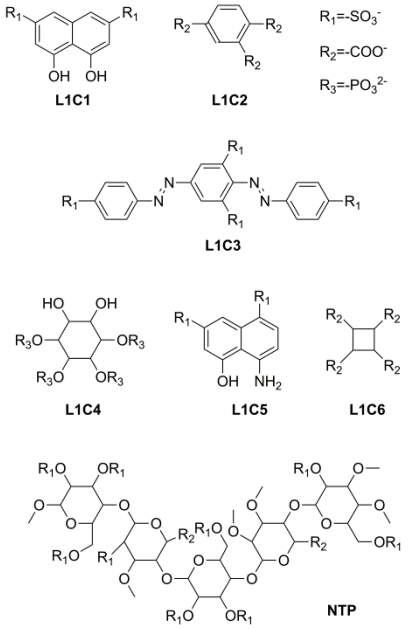

Fig. 1. Compounds for antithrombin target. Compound NTP is synthetic pentasaccharide compounds from crystal structure (PDB ID: 1AZX).

In this study, we have gathered a panel of antithrombin ligands (Figure 1), including heparin and non-polysaccharide scaffold compounds. Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) method is employed in the binding affinity calculation of antithrombin ligands. MM/GBSA method is selected because it is the fastest force-field based method that computes the free energy of binding, as compared to the other free energy computation methods, such as free energy perturbation (FEP) and thermodynamic integration (TI) methods[8]. MM/GBSA method has been widely exploited in free energy calculations[9,10]. In most scenarios, the entropy term is neglected in the calculation for relative free binding energies. Quite a few researchers dispute the benefits of including entropy term, which can be a major source of error due to the drawback of the entropy calculation method[11,12], despite others advocate its usage[13]. In this study, we neglect the entropy calculations.

2 Method

The MM/GBSA calculations are applied to the antithrombin (PDB ID: 1AZX) and its 7 ligands (Figure 1). The Amber forcefield f99SB[14] is employed in the calculation for the antithrombin receptor. Ligands use the Amber GAFF forcefield[15] as determined

by the antechamber program[16] in the Amber package. Partial charges of ligands are calculated using the AM1-BCC method.[17] The receptor-ligand complexes are energetically minimized by the MM/GBSA method implemented in the Sander program of the Amber package[18]. The atomic radii developed by Onufriev and coworkers (Amber input parameter igb=5) are chosen for all GB calculations[19]. The systems are heated from 0 K to room temperature, 300 K. The MD simulations with a time step of 2 fs for the integration of the equations of motion were carried out at room temperature. The systems are equilibrated at room temperature for 500 ps. Each MD trajectory is followed to 100 ns after equilibrium. Binding affinities of antithrombin and its 7 ligands are calculated by post-processing the ensembles of structures extracted from MD trajectories using MM/GBSA calculations. In the MM/GBSA calculation, the binding free energy between a receptor and a ligand is calculated as following equations:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{receptor}} - G_{\text{ligand}}$$
$$\Delta G_{\text{bind}} = \Delta H - T\Delta S \approx \Delta E_{\text{gas}} + \Delta G_{\text{sol}} - T\Delta S$$
$$\Delta E_{\text{gas}} = \Delta E_{\text{int}} + \Delta E_{\text{ELE}} + \Delta E_{\text{VDW}}$$
$$\Delta G_{\text{Sol}} = \Delta G_{\text{GB}} + \Delta G_{\text{Surf}}$$

The binding free energy (ΔG_{bind}) is decomposed into different energy terms. Because the structures of complex, receptor, and ligand are extracted from the same trajectory, the internal energy change (ΔE_{int}) is canceled. Thus, the gas-phase interaction energy (ΔE_{gas}) between the receptor and the ligand is the sum of electrostatic (ΔE_{ELE}) and van der Waals (ΔE_{VDW}) interaction energies. The solvation free energy (ΔG_{Sol}) is divided into the polar and non-polar energy terms. The polar solvation energy (ΔG_{GB}) is calculated by using GB model. The non-polar contribution is calculated based on the solvent-accessible surface area (ΔG_{Surf}). A value of 80 is used for the solvent dielectric constant and the solute dielectric constant is set to 1. The calculated binding free energy (ΔG_{bind}) is the sum of the gas-phase interaction energy and solvation free energy because we neglect the entropy term. The experimental binding free energy is estimated from the experimental dissociation constant (Kd) by the equation:

$$\Delta G_{\text{Exp}} = \text{RT} \cdot \ln(\text{Kd})$$

where R is the gas constant and T is the temperature.

3 Results and Discussion

The calculated binding free energies of 7 antithrombin ligands are shown in the Table 1 together with their corresponding experimental values. Each calculated binding free energy is averaged from snapshots extracted from 100 ns MD trajectory. Except for Compound L1C1, all the antithrombin ligands have experimental binding free energies. Determined by the experiments, Compound L1C4 is the best binder with a Kd value of 0.088 uM. As predicted by the MM/GBSA method, Compound L1C4 has the

most negative binding free energy (-308.01 kcal/mol), which is in agreement with the experimental results. The second best binder, Compound NTP, is ranked in second place in the MM/GBSA calculation with a calculated binding free energy of -279.57 kcal/mol. This prediction also agrees with the experimental data. Compound L1C2 is predicted to have the worst binding free energy in 6 ligands, which is in agreement with its experimental value. In summary, the MM/GBSA calculations rank the binding affinities of 6 antithrombin ligands in same order as that of experimental binding free energies. The calculated binding free energies of 6 antithrombin ligands have been plotted against the experimental ones. The correlation coefficient (R2) is 0.69, which indicates good correlation between the calculated and experimental values.

Table 1. Calculated and experimental binding free energies (kcal/mol) of antithrombin ligands.

Cmpd	ΔE_{ELE}	ΔE_{VDW}	ΔE_{gas}	ΔG_{Surf}	ΔG_{GB}	$\Delta G_{GB\text{-}ELE}$	ΔG_{Sol}	ΔG_{bind}	Kd(uM)	ΔG_{Exp}
L1C1	-552.67	-23.68	-576.35	-2.75	480.12	-72.55	477.37	-98.97	-	-
L1C2	-442.99	0.47	-442.52	-1.20	417.60	-25.39	416.41	-26.11	13700	-2.54
L1C3	-836.77	-39.96	-876.73	-4.06	781.94	-54.83	777.88	-98.85	10.02	-6.81
L1C4	-1599.09	33.02	-1566.07	-2.98	1261.05	-338.04	1258.07	-308.01	0.088	-9.62
L1C5	-613.30	-19.00	-632.31	-2.57	525.82	-87.48	523.25	-109.06	0.69	-8.40
L1C6	-818.73	8.21	-810.52	-1.54	752.94	-65.79	751.41	-59.11	17.52	-6.48
NTP	-2598.87	-60.89	-2659.76	-7.58	2387.77	-211.09	2380.20	-279.57	0.104	-9.52

As shown in the figure 1, all antithrombin ligands contain negatively charged groups, which suggests electrostatic interaction should be a key factor for the binding affinity. Compound NTP has a total charge of -11 and Compound L1C4 has a total charge of -8. By decomposition of the binding free energy, Compound NTP and L1C4 have largest electrostatic energy changes upon binding in both gas phase (ΔE_{ELE}) and GB solvent ($\Delta G_{GB\text{-}ELE}$). Energy change in gas phase is actually the energy change for the solute. Thus, in other words, Compound NTP and L1C4 have largest electrostatic energy changes upon binding in solute and solvent. In contrast, Compound L1C2 has the smallest electrostatic energy changes in solute and solvent. Although Compound L1C4 has the least favor of van der Waals energy change upon binding, it has been compensated by the significant electrostatic energy changes. For all ligands, the van der Waals energy changes (ΔE_{VDW}) upon binding are less than the electrostatic energy changes (ΔE_{ELE}) by 1-2 orders magnitudes. The contribution of the van der Waals energy change has been overpowered by the electrostatic energy change. The sizes of Compound NTP and L1C3 are larger than that of the rest compounds. Thus, non-polar contribution of solvation free energy of the Compound NTP and L1C3 are more negative than that of the rest compounds. Nevertheless, non-polar contributions for compounds are all small. Non-polar contribution is overwhelmed by the polar contribution of solvation free energy. The two major factors to determine the binding affinity are electrostatic energy change and solvation free energy change. The higher the total charge of compound, the higher penalty it has to pay for solvation free energy. However, high penalty for high total charge of compound has been paid off by favorable electrostatic energy changes. Although the electrostatic energy change of Compound L1C4 is less than that of Compound NTP, Compound L1C4 needs less compensation

for the solvation free energy. Thus, Compound L1C4 is better binder than Compound NTP.

Hydrogen bonding analysis has been carried out for MD trajectories using ptraj program in the Amber package. For Compound NTP and L1C4, the numbers of hydrogen bonds to antithrombin with the time occupy larger than 20% are 40 and 25, respectively. For Compound L1C1, L1C2, L1C3, L1C5, and L1C6, that numbers are 5, 10, 12, 12, and 12 respectively. Taking the molecular weight into account and using a similar approach as Reynolds' ligand efficiency method[20], Compound L1C4 has the highest ligand efficiency.

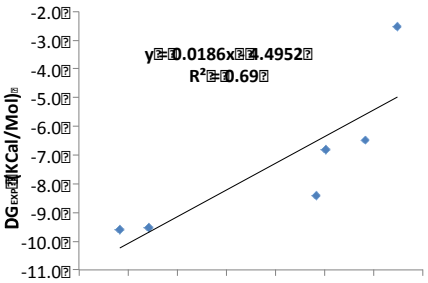

Fig. 2. The scatter plot of calculated MM/GBSA binding free energy versus experimental binding affinity estimated from dissociation constant

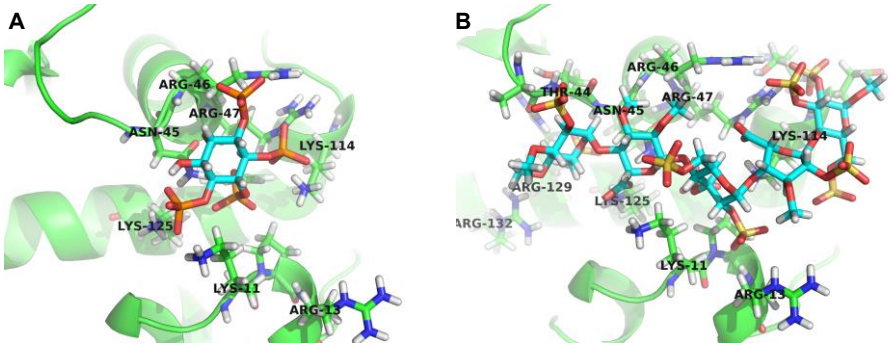

Fig. 3. Initial structures of Compound L1C4 (A) and NTP (B) complex with antithrombin.

Compound L1C4 forms double hydrogen bonds with Arg47 (Figure 3A). One hydrogen bond (O6-HH21-NH2) has 94.81% time occupancy and the other one (O6-HE-NE) has 89.55%. The average hydrogen bond distances between the heavy atoms are 2.74 Å and 2.70 Å, respectively. In a nutshell, Compound L1C4 has strong hydrogen bonds with Arg47 and one of the four phosphate groups from Compound L1C4 is locked to the Arg47. According to the hydrogen bonding analysis, Compound L1C4 is also hydrogen bond to Arg46, Arg13, Lys114, Lys11, Lys125, and Asn45, which are key residues to the binding process. We found that the binding of Compound L1C4 to antithrombin is non-specific. Except for the one locked to Arg47, the rest three phosphate groups of Compound L1C4 can rotate so that key residues can form hydrogen bond to different oxygen atoms of phosphate at different time period of MD trajectory. It is worth noting that Arg13 is far away from the Compound L1C4 at the starting conformation. After 8 ns of MD simulation, Arg13 begins to make hydrogen bonds with phosphate group of Compound L1C4, which suggests that long-time MD simulation is essential to obtained accurate binding affinity. As shown in Figure 3B,

Compound NTP makes hydrogen bond to antithrombin mainly via its negatively charged sulfate groups. Compound NTP forms hydrogen bonds with Arg13, Arg129, Arg47, and Asn45 with high time occupancy (70~88%). It forms hydrogen bonds with Arg132, Lys125, Thr44 with median time occupancy (43~66%). It also forms relatively weak hydrogen bonds with Arg46, LYS114 and LYS11.

Judge from the hydrogen bond analysis on Compound L1C4 and NTP, Arg47, Arg13, and Asn45 play crucial role in the antithrombin binding process. Antithrombin provides multiple sulfate/phosphate binding sites consisted of mostly positively charged residues (arginine, lysine) and neutral charged residues that can provide rich hydrogen bond donors/acceptors (asparagine). All four phosphate groups of Compound L1C4 form hydrogen bonds with antithrombin while not all sulfate groups of Compound NTP can form hydrogen bonds with antithrombin. As pointed out above, introducing positively charged group in the ligand will result in the penalty in solvation free energy. If adding a positively charged group cannot form favorable interactions (e.g. hydrogen bonding), ligand efficiency will be reduced. This can explain that Compound L1C4 has higher ligand efficiency than Compound NTP.

4 Conclusion

In this study, we calculated the binding affinities of 7 antithrombin ligands by employing the MM/GBSA method. The plot of calculated binding free energies versus experimental ones has shown good correlation. The electrostatic interactions in both solute and solvent contribute favorably to the binding free energy. Adding more negatively charged groups to ligand provides more favorable electrostatic energy change. However, it creates higher penalty for the solvation free energy simultaneously. The penalty can be compensated by forming more hydrogen bonds as more negatively charged groups are added into ligand. Compound L1C4 has higher ligand efficiency because it uses all its phosphate groups to form hydrogen bonds with antithrombin while Compound NTP does not.

5 Acknowledgement

The authors thank Livermore Computing for the computer time and Laboratory Directed Research and Development for funding (12-SI-004). We also thank Livermore Computing Grand Challenge for extensive computing resources. This work was performed under the auspices of the United States Department of Energy by the Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. Release Number (To be inserted).

6 Reference:

(1) Turko, I. V.; Fan, B.; Gettins, P. G. W. *FEBS Lett.*, *335*, 9.

(2) Jin, L.; Abrahams, J. P.; Skinner, R.; Petitou, M.; Pike, R. N.; Carrell, R. W. *Proceedings of the National Academy of Sciences* **1997**, *94*, 14683.
(3) Thunberg, L.; Bäckström, G.; Lindahl, U. *Carbohydr. Res.* **1982**, *100*, 393.
(4) Verli, H.; Guimarães, J. A. *J. Mol. Graph. Model.* **2005**, *24*, 203.
(5) Pol-Fachin, L.; Franco Becker, C.; Almeida Guimarães, J.; Verli, H. *Proteins: Structure, Function, and Bioinformatics* **2011**, *79*, 2735.
(6) Bitomsky, W.; Wade, R. C. *J. Am. Chem. Soc.* **1999**, *121*, 3004.
(7) Navarro-Fernández, J.; Pérez-Sánchez, H.; Martínez-Martínez, I.; Meliciani, I.; Guerrero, J. A.; Vicente, V.; Corral, J.; Wenzel, W. *J. Med. Chem.* **2012**, *55*, 6403.
(8) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431.
(9) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. *Acc. Chem. Res.* **2000**, *33*, 889.
(10) Hou, T. J.; Wang, J. M.; Li, Y. Y.; Wang, W. *J. Chem. Inf. Model.* **2011**, *51*, 69.
(11) Rastelli, G.; Rio, A. D.; Degliesposti, G.; Sgobba, M. *J. Comput. Chem.* **2010**, *31*, 797.
(12) Greenidge, P. A.; Kramer, C.; Mozziconacci, J. C.; Wolf, R. M. *J. Chem. Inf. Model.* **2013**, *53*, 201.
(13) Lafont, V.; Armstrong, A. A.; Ohtaka, H.; Kiso, Y.; Mario Amzel, L.; Freire, E. *Chemical Biology & Drug Design* **2007**, *69*, 413.
(14) Ponder, J. W.; Case, D. A. *Protein Simulations* **2003**, *66*, 27.
(15) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.
(16) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graph. Model.* **2006**, *25*, 247.
(17) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2000**, *21*, 132.
(18) Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668.
(19) Onufriev, A.; Bashford, D.; Case, D. A. *Proteins-Struct. Funct. Bioinf.* **2004**, *55*, 383.
(20) Reynolds, C. H.; Tounge, B. A.; Bembenek, S. D. *J. Med. Chem.* **2008**, *51*, 2432.